# Take home exam
# Part II

## *Convex Unconstrained and Constrained Optimization*

March 27, 2022

Gloria del Valle Cano
gloria.valle@estudiante.uam.es

---

**Problem 1.** (1 point) We have worked out the elementary version of Lagrange multipliers assuming that from $g(x, y) = 0$ we can find a function $y = h(x)$ such that $g(x, h(x)) = 0$.
But sometimes what we get is that there is an $h$ such that $g(h(y), y) = 0$. Rewrite the Lagrange multiplier analysis in the lecture slides under this assumption.

For $f, g : \mathbb{R}^2 \to \mathbb{R}$ consider the following minimization problem

$$\min f(x, y) \text{ s.t. } g(x, y) = 0.$$

Assuming the **Implicit Function Theorem** holds, we can find a function $x = h(y)$ s.t. $g(h(y), y) = 0$ and, thus, we can write

$$f(x, y) = f(h(y), y) = \Psi(y).$$

At a minimum $y^*$ with $x^* = h(y^*)$ we thus have

$$0 = \Psi'(y^*) = \frac{\partial g}{\partial x}(x^*, y^*)h'(y^*) + \frac{\partial g}{\partial y}(x^*, y^*). \tag{1}$$

But since $g(h(y), y) = 0$, we also have

$$0 = \frac{\partial g}{\partial x}(x^*, y^*)h'(y^*) + \frac{\partial g}{\partial y}(x^*, y^*) \implies h'(y^*) = -\frac{\frac{\partial g}{\partial y}(x^*, y^*)}{\frac{\partial g}{\partial x}(x^*, y^*)}. \tag{2}$$

Putting together 1 and 2 we arrive at

$$0 = \frac{\partial f}{\partial y}(x^*, y^*)\frac{\partial g}{\partial x}(x^*, y^*) - \frac{\partial f}{\partial x}(x^*, y^*)\frac{\partial g}{\partial y}(x^*, y^*).$$

That is, at $(x^*, y^*)$, $\nabla f \perp \left(-\frac{\partial g}{\partial y}, \frac{\partial g}{\partial x}\right)$ and, since $\left(-\frac{\partial g}{\partial y}, \frac{\partial g}{\partial x}\right) \perp \nabla g$, we have $\nabla f \| \nabla g$ i.e. $\nabla f(x^*, y^*) = -\lambda^* \nabla g(x^*, y^*)$ for some $\lambda^* \neq 0$.

Thus, for the **Lagrangian**

$$\mathcal{L}(x, y, \lambda) = f(x, y) + \lambda g(x, y),$$

we have a minimum $(x^*, y^*)$ there is a $\lambda^* \neq 0$ s.t.

$$\nabla \mathcal{L}(x^*, y^*, \lambda^*) = \nabla f(x^*, y^*) + \lambda^* \nabla g(x^*, y^*) = 0.$$

**Problem 2.** (3 points) We want to solve the following constrained minimization problem:

$$\begin{aligned} \min \quad & f(x,y) = x^2 + 2xy + 2y^2 - 3x + y \\ \text{s.t.} \quad & x + y = 1, \\ & x \geq 0, y \geq 0. \end{aligned}$$

Argue first that $f$ is convex and then:

- Write its Lagrangian with $\alpha, \beta$ the multipliers of the inequality constraints.
- Write the KKT conditions.
- Use them to solve the problem. For this consider separately the $(\alpha = \beta = 0), (\alpha > 0, \beta = 0), (\alpha = 0, \beta > 0), (\alpha > 0, \beta > 0)$ cases.

First of all, we verify that $f$ is convex because its Hessian matrix is positive semidefinite, or equivalently its eigenvalues are non-negative. The Hessian matrix is

$$\mathbf{H}_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 4 \end{bmatrix},$$

whose eigenvalues are $\lambda_1 = 3 + \sqrt{5}$ and $\lambda_2 = 3 - \sqrt{5}$, both positive. As a result, we determine $f$ is convex.

Now we write its Lagrangian, with $\alpha$ and $\beta$ as multipliers of the inequality constraints and $\lambda$ as equality constraint multiplier.

$$\mathcal{L}(x, y, \lambda, \alpha, \beta) = x^2 + 2xy + 2y^2 - 3x + y + \lambda(x + y - 1) - \alpha x - \beta y.$$

Assuming that the hypothesis of the KKT conditions theorem hold, the resulting KKT conditions on a local minimum $(x^*, y^*)$ are the following:

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x}(x^*, y^*, \lambda, \alpha, \beta) \implies \lambda + 2x^* + 2y^* - 3 - \alpha = 0, \\ 0 &= \frac{\partial \mathcal{L}}{\partial y}(x^*, y^*, \lambda, \alpha, \beta) \implies \lambda + 2x^* + 4y^* + 1 - \beta = 0, \\ 0 &= \alpha x^*, \\ 0 &= \beta y^*. \end{aligned}$$

Then, we use them to solve the problem, considering the four possible cases below:

- Case $\alpha = \beta = 0$.

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial x} = \lambda + 2x + 2y - 3 \\ 0 &= \frac{\partial \mathcal{L}}{\partial y} = \lambda + 2x + 4y + 1 \end{aligned}$$

From both expressions we get $2x + 2y - 3 = 2x + 4y + 1 \implies 2y = -4 \implies y = -2$. Then, $x = 1 - y = 3$ and $\lambda = -1$. So we have $(\mathbf{3}, -\mathbf{2})'$ as **feasible KKT point**.

- Case $\alpha > 0, \beta = 0$. When $\alpha > 0$, $x = 0$, so $y = 1$ and $\lambda = -5$. Therefore we have a **feasible KKT point** on $(\mathbf{0}, \mathbf{1})'$.

- Case $\alpha = 0, \beta > 0$. When $\beta > 0$, $y = 0$, so $x = 1$ and $\lambda = 1$. Therefore we have a **feasible KKT point** on $(\mathbf{1}, \mathbf{0})'$.

- Case $\alpha > 0, \beta > 0$. This implies $x = y = 0$, so we have a contradiction because $x + y \neq 1$. This means that $(0, 0)'$ **is not** a feasible KKT point.

Given all points found, we can determine that our optimal solution is $(\mathbf{1}, \mathbf{0})'$, with an optimal value $\lambda = \mathbf{1}$.

**Problem 3.** (1 point) Let $f : S \subset \mathbb{R}^d \to \mathbb{R}$ be a convex function on the convex set $S$ and we extend it to an $\tilde{f} : \mathbb{R}^d \to \mathbb{R}$ as:

$$\begin{aligned} \tilde{f}(x) &= f(x) \text{ if } x \in S. \\ &= +\infty \text{ if } x \notin S. \end{aligned}$$

Show that $\tilde{f}$ is a convex function on $\mathbb{R}^d$. Assume that $a + \infty = \infty$ and that $a \cdot \infty = \infty$ for $a > 0$.

We say that $S$ is a **convex set** if for all $x, x' \in S$ and $\lambda \in [0, 1]$,

$$\lambda x + (1 - \lambda)x' \in S.$$

Let $x, x' \in S$ and $\lambda \in [0, 1]$, so here we cover two cases:

- *First case.* If $x, x' \in S$:

  $$\tilde{f}(\lambda x + (1 - \lambda)x') = f(\lambda x + (1 - \lambda)x') \leq \lambda f(x) + (1 - \lambda)f(x') = \lambda\tilde{f}(x) + (1 - \lambda)\tilde{f}(x').$$

  Where the first equality holds given that $S$ is convex. The inequality holds because $f$ is convex. And, the last equality raises from the definition of $\tilde{f}$.

- *Second case.* If $x \notin S$ or $x' \notin S$, we have that

  $$\tilde{f}(\lambda x + (1 - \lambda)x') \leq \lambda\tilde{f}(x) + (1 - \lambda)\tilde{f}(x') = +\infty,$$

  because $\tilde{f}(y) \leq +\infty$, $\forall y \in \mathbb{R}^d$.

Since both cases satisfy convexity definition holds, we conclude this function is convex.

**Problem 4.** (2 points) Prove **Jensen's inequality**: if $f$ is convex on $\mathbb{R}^d$ and $\sum_1^k \lambda_i = 1$, with $0 \le \lambda_i \le 1$ we have for any $x_1, \dots, x_k \in \mathbb{R}^d$

$$f\left(\sum_1^k \lambda_i x_i\right) \le \sum_1^k \lambda_i f(x_i)$$

*Hint: just write* $\sum_1^k \lambda_i x_i = \lambda_1 x_1 + (1-\lambda_1)$ *for an appropriate* $v$ *and apply repeatedly the definition of a convex function. Start with* $k=3$ *and carry on.*

We proceed using an inductive procedure:

- If $k = 1$ then $\lambda = 1$, so we simply have $f(x_1) = f(x_1)$, which is true, and nothing to prove. If $k = 2$ we have the definition of the convexity of $f$:

$$\lambda_1 + \lambda_2 = 1,\ \ \lambda_1, \lambda_2 \ge 0 \implies f(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

- We assume the statement its true for $k$ and consider $k+1$ points $x_1, \dots, x_{k+1}$, with coefficients $\lambda_1, \dots, \lambda_{k+1} \ge 0$, $\sum_{i=1}^{k+1} \lambda_i = 1$. The evaluation of the linear combination can be decomposed as

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) = f\left((1-\lambda_1)\left(\sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} x_i\right) + \lambda_1 x_1\right).$$

  Using this, it is straightforward to use the Jensen's inequality on $x_1$ and $\sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} x_i$ with coefficients $\lambda_1$ and $1-\lambda_1$ respectively. That is,

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) \le (1-\lambda_1) f\left(\sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} x_i\right) + \lambda_1 f(x_1).$$

  We may notice that $\sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} x_i$ verifies the inductive hypothesis, thus,

$$f\left(\sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} x_i\right) \le \sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} f(x_i).$$

  Finally,

$$f\left(\sum_{i=1}^{k+1} \lambda_i x_i\right) \le (1-\lambda_1) \sum_{i=2}^{k+1} \frac{\lambda_i}{1-\lambda_1} f(x_i) + \lambda_1 f(x_1) = \sum_{i=1}^{k+1} \lambda_i f(x_i).$$

**Problem 5.** (3 points) Prove that the following function is convex

$$\begin{aligned} f(x) &= x^2 - 1, \qquad && |x| > 1 \\ &= 0 && |x| \le 1 \end{aligned}$$

and compute its proximal. Which are the fixed points of this proximal?